A priori tensor approximations for the numerical solution of high dimensional problems: alternative definitions

Anthony Nouy

Proper Generalized Decompositions for a priori model reduction of problems formulated in tensor product spaces: alternative definitions and algorithms

A. Nouy*

* *GeM - UMR CNRS 6183, Ecole Centrale Nantes, Université de Nantes*
1 rue de la Noë, 44321 Nantes, France
anthony.nouy@ec-nantes.fr

Abstract

Model reduction techniques based on the construction of separated representations are receiving a growing interest in scientific computing. A family of methods, recently called Proper Generalized Decomposition (PGD) methods, have been introduced for the a priori construction of separated representations of the solution of problems defined in tensor product spaces. Different definitions of PGDs and associated algorithms have been proposed. Here, we review classical definitions of progressive PGDs and we introduce several variants in order to improve convergence properties. In particular, we introduce the Minimax PGD, recently proposed for evolution problems, which can be interpreted as a Petrov-Galerkin model reduction technique. The different variants are presented in an abstract setting. Model examples will illustrate some properties of the different variants of PGDs.

Keywords: Proper Generalized Decomposition (PGD), Model reduction, Separated representations, Tensor product approximation, Galerkin PGD, Minimal Residual PGD, Minimax PGD

1 Introduction

Model reduction techniques based on the construction of separated representations are receiving a growing interest in scientific computing. A family of methods, re-

cently called Proper Generalized Decomposition (PGD) methods [1, 2, 3], have been introduced for the a priori construction of separated representations of the solution u of problems defined in tensor product spaces:

$$u \in V = V_1 \otimes \ldots \otimes V_d, \quad A(u) = l$$

PGD methods can be interpreted as generalizations of Proper Orthogonal Decomposition (or Singular Value Decomposition, or Karhunen-Loève Decomposition) for the a priori construction of a separated representation u_m of the solution u (i.e. without knowing u a priori):

$$u \approx u_m = \sum_{i=1}^{m} w_i^1 \otimes \ldots \otimes w_i^d, \quad w_i^k \in V_k \tag{1}$$

Several definitions of PGDs have been proposed. Basic PGDs are based on a progressive construction of the sequence u_m, where at each step, an additional rank-one element $\otimes_{k=1}^{d} w_m^k$ is added to the previously computed decomposition u_{m-1} [4, 5, 6]. These progressive definitions of PGDs can then be considered as Greedy algorithms for constructing separated representations (analogy introduced in [7]). The additional rank-one element is classically defined by using Galerkin or Minimal Residual criteria. On one hand, the progressive PGD based on a minimal residual formulation is robust in the sense that the sequence u_m monotically converges with respect to the residual norm. However, the convergence with respect to useful norms can be very slow. On another hand, the progressive PGD based on Galerkin orthogonality is not robust since monotonic convergence is not guaranteed for general non-symmetric problems. Moreover, the sequence u_m may diverge in some situations. However, when convergent, Galerkin PGD should be preferred since it relies on classical formulations of problems, it requires less computational efforts and it classically leads to better convergence properties with respect to standard metrics. A possible improvement of these progressive decompositions consists in introducing some updating steps in order to capture an approximation of the optimal decomposition [8, 3], obtained by defining the whole set of functions simultaneously (and not progressively). For many applications, it allows recovering good convergence properties of separated representations. However, for some large scale applications, these updating steps may be computationally expansive or simply unaffordable. A new definition of PGD, called Minimax PGD, has been recently proposed in the context of time-dependent partial differential equations. It can be interpreted as a Petrov-Galerkin model reduction technique, where "test' and "trial" reduced basis functions are related by an adjoint problem involving a predefined scalar product on Hilbert space V. Knowing an approximation u_{m-1}, an additional rank-one element $\otimes_{k=1}^{d} w_m^k$ (the trial function) is computed along with another rank-one element $\otimes_{k=1}^{d} \tilde{w}_m^k$ (the test function). These two elements are defined as the solutions of a saddle-point problem. This new definition can significantly improve convergence properties of progressive separated representations with respect to the norm associated with the chosen scalar product.

In this paper, we present the above alternative definitions for general problems formulated in tensor product Hilbert spaces. The abstract setting is presented in section 2. Section 3 presents different definitions of progressive PGDs, based on a progressive construction of decomposition (1). In section 5, the behavior of the different variants of PGDs is illustrated on two model examples: separation of spatial coordinates for partial differential equations defined on hyper-rectangular domains, separation of physical variables and random parameters in stochastic partial differential equations.

2 Problems defined in tensor product spaces

2.1 Abstract formulation

We consider an abstract formulation of a problem defined in a Hilbert space V:

$$u \in V, \quad \mathcal{A}(u,v) = \mathcal{L}(v) \quad \forall v \in V \tag{2}$$

where $\mathcal{A}$ is a bilinear or eventually semilinear form on V, and $\mathcal{L}$ is a linear form on V. We consider $V = V_1 \otimes \ldots \otimes V_d$ as a tensor product of Hilbert spaces V_k. We introduce on V an inner product $(\cdot,\cdot)$ and associated norm $\|\cdot\|$, constructed from inner products $(\cdot,\cdot)_i$ and norms $\|\cdot\|_i$ on Hilbert spaces V_k:

$$(\otimes_{k=1}^d w_k, \otimes_{k=1}^d \tilde{w}_k) = \prod_{k=1}^d (w_k, \tilde{w}_k)_k, \quad \| \otimes_{k=1}^d w_k \| = \prod_{k=1}^d \|w_k\|_k \tag{3}$$

We introduce the operator $A : V \to V$ and element $l \in V$ associated with $\mathcal{A}$ and $\mathcal{L}$ respectively, defined by

$$\mathcal{A}(u,v) = (v, Au), \quad \mathcal{L}(v) = (v, l),$$

for all $u, v \in V$.

2.2 Separated representations

We introduce the set of rank-one tensors

$$\mathcal{S}_1 = \{z = w^1 \otimes \ldots \otimes w^d \ ; \ w^k \in V_k, \ k \in \{1 \ldots d\}\}, \tag{4}$$

and the set of rank-m tensors

$$\mathcal{S}_m = \{v = \sum_{i=1}^m z_i \ ; \ z_i \in \mathcal{S}_1, \ i \in \{1 \ldots m\}\} \tag{5}$$

The construction of a separated representation $u_m \in \mathcal{S}_m$ of a given element $u \in V$ has been extensively studied over the past years in multilinear algebra community and different definitions and associated algorithms have been proposed [9, 10, 11, 12]. These

definitions can be seen as multidimensional versions of the Singular Value Decomposition (SVD). The question of finding an optimal decomposition of a given rank m is not a trivial question and has led to various definitions of separated representations. An optimal representation $u_m \in \mathcal{S}_m$ can not be simply defined as an optimization problem on $\mathcal{S}_m$ since it may lead to an ill-posed problem (for $d \geq 3$). In fact, suitable constraints (orthogonality or boundedness constraints) must be imposed in order to define an optimization problem in a suitable subset of $\mathcal{S}_m$ (see section 4).

In this paper, we address the more difficult problem of finding a separated representation of the solution u of problem (2), without computing the solution u *a priori*, which is simply impossible for high dimensional problems. It is the aim of the Proper Generalized Decomposition (PGD) method, which can be seen as a generalization of a multidimensional SVD. In the following sections, we propose and compare different possible definitions of PGD.

2.3 Model examples

2.3.1 Example 1

As a first example, we consider a linear elliptic partial differential equation on a hyper-rectangular domain $\Omega = \Omega_1 \times \ldots \times \Omega_d$, with $V_k = H_0^1(\Omega_k)$, and

$$\mathcal{A}(u, v) = \int_\Omega \nu \nabla u \cdot \nabla v d\mu + \int_\Omega v\, c \cdot \nabla u d\mu \tag{6}$$

$$\mathcal{L}(v) = \int_\Omega f v d\mu \tag{7}$$

where $d\mu = dx_1 \ldots dx_d$, $f \in L^2(\Omega)$, $c \in L^\infty(\Omega)^d$, $0 < \nu, \nu^{-1} \in L^\infty(\Omega)$. A separated representation $u_m \in \mathcal{S}_m$ reads

$$u_m(x) \equiv \sum_{i=1}^{m} w_i^1(x_1) \ldots w_i^d(x_d)$$

2.3.2 Example 2

As a second example, we consider a parameterized partial differential equation defined on a domain $\Omega = \times_{k=1}^n \Omega_k$ and a parameter space $\Xi = \times_{k=1}^s \Xi_k \subset \mathbb{R}^s$. It means that the operator and right-hand side depend on a finite set of parameters $\xi \in \Xi$. The parameter space is endowed with a product probability measure $P_\xi = \otimes_{k=1}^s P_{\xi_k}$. The parameters $\xi = (\xi_1, \ldots, \xi_s)$ are then considered as independent real-valued random variables. We let $V_k \subset H^1(\Omega_k)$ for $k \in \{1 \ldots n\}$, and $V_{n+k} = L^2_{P_{\xi_k}}(\Xi_k)$ for $k \in \{1 \ldots s\}$. We then

let

$$\mathcal{A}(u,v) = \int_{\Omega\times\Xi} (\nu(x,y)\nabla u\cdot\nabla v + v\, c(x,y)\cdot\nabla u + \sigma(x,y)u\, v)\, d\mu(x,y) \tag{8}$$

$$\mathcal{L}(v) = \int_{\Omega\times\Xi} v\, f(x,y)\, d\mu(x,y) \tag{9}$$

where $d\mu(x,y) = dxdP_\xi(y) = dx_1\dots dx_n dP_{\xi_1}(y_1)\dots dP_{\xi_s}(y_s)$, $f \in L^2(\Omega\times\Xi)$, $c \in L^\infty(\Omega\times\Xi)^n$, $0 < \nu, \nu^{-1} \in L^\infty(\Omega\times\Xi)$, $0 \le \sigma \in L^\infty(\Omega\times\Xi)$. A separated representation $u_m \in \mathcal{S}_m$ reads

$$u_m(x,\xi) \equiv \sum_{i=1}^{m} w_i^1(x_1)\dots w_i^n(x_n) w_i^{n+1}(\xi_1)\dots w_i^{n+s}(\xi_s)$$

3 Progressive Proper Generalized Decomposition

In this section, we present different definitions of PGDs for some classes of linear and nonlinear problems. We here focus on progressive PGDs, based on a progressive construction of the separated representation

$$u_m = \sum_{i=1}^{m} z_i, \quad z_i \in \mathcal{S}_1 \tag{10}$$

We suppose that a rank-m decomposition $u_m \in \mathcal{S}_m$ has been determined. The aim is then to define a new function $z_{m+1} \in \mathcal{S}_1$, leading to the following rank-$(m+1)$ decomposition:

$$u_{m+1} = u_m + z_{m+1}, \quad z_{m+1} \in \mathcal{S}_1 \tag{11}$$

This progressive construction can be interpreted as a Greedy algorithm for constructing a separated representation. We begin in section 3.1 with the well-established case of problems associated with convex optimization problems, for which there is a natural definition of the progressive PGD. Then, in section 3.2, we present different alternative definitions for more general cases.

3.1 Variational problems associated with convex optimization

We first consider the particular case where equation (2) is the Euler-Lagrange equation of a convex optimization problem, i.e. such that

$$\mathcal{A}(u,v) - \mathcal{L}(v) = (\mathcal{J}'(u), v)$$

where $\mathcal{J}' : V \to V$ is the differential of a strongly convex, coercive and differentiable functional $\mathcal{J} : V \to \mathbb{R}$. Equation (2) is then equivalent to the following minimization problem:

$$\mathcal{J}(u) = \min_{v\in V} \mathcal{J}(v) \tag{12}$$

3.1.1 Definition of progressive PGD.

The progressive PGD (10) is naturally defined as follows:

$$\mathcal{J}(u_m + z_{m+1}) = \min_{z \in \mathcal{S}_1} \mathcal{J}(u_m + z) \tag{13}$$

The existence of a minimizer follows from the properties of $\mathcal{J}$ (convexity, continuity, coercivity) and from the fact that $\mathcal{S}_1$ is weakly closed in V [13]. For general nonlinear convex problems, convergence results on the sequence u_m can be found in [14].

3.1.2 Linear case: interpretation as a generalized multidimensional SVD.

For the particular case of linear elliptic problems, where $\mathcal{A}$ is a symmetric continuous coercive bilinear form on V, we classically have

$$\mathcal{J}(v) = \frac{1}{2}\mathcal{A}(v, v) - \mathcal{L}(v) \tag{14}$$

A proof for the strong convergence of u_m to u can be found in [13]. In this particular case, the PGD can be interpreted as a generalization of the Singular Value Decomposition (SVD) in the following sense. $\mathcal{A}$ defines a norm $\|\cdot\|_{\mathcal{A}} : v \mapsto \mathcal{A}(v, v)^{1/2}$, which is equivalent to usual norm $\|\cdot\|$ defined in (3). Then, definition (13) is equivalent to

$$\|u - u_m - z_{m+1}\|_{\mathcal{A}}^2 = \min_{z \in \mathcal{S}_1} \|u - u_m - z\|_{\mathcal{A}}^2 \tag{15}$$

and therefore, z_{m+1} is the optimal rank-one correction of u_m with respect to the norm induced by $\mathcal{A}$. We have

$$\|u - u_m\|_{\mathcal{A}}^2 = \|u - u_{m-1}\|_{\mathcal{A}}^2 - \sigma_m^2 = \|u\|_{\mathcal{A}}^2 - \sum_{i=1}^{m} \sigma_i^2 \underset{m \to \infty}{\longrightarrow} 0 \tag{16}$$

where $\sigma_m = \|z_m\|_{\mathcal{A}}$ can be interpreted as the dominant singular value of $(u - u_m)$ associated with the metric induced by $\mathcal{A}$, characterized by

$$\sigma_m = \max_{w \in \mathcal{S}_1; \|w\|_{\mathcal{A}} = 1} (u - u_{m-1}, w)_{\mathcal{A}}$$

In this context, the progressive PGD appears as a generalization of a multidimensional SVD, which satisfies an optimality property with respect to the metric induced by the operator of a linear symmetric elliptic problem.

3.1.3 Alternated direction algorithm.

A possible way for constructing $z_{m+1} \in \mathcal{S}_1$ consists in using an alternated direction algorithm, by solving successively problems of type

$$\min_{w^l \in V_l} \mathcal{J}(u_m + \otimes_{k=1}^{d} w^k) \tag{17}$$

For a given $z = \otimes_{k=1}^{d} w^k$, we introduce the linear subspace $\mathcal{S}_1^l(z) \subset \mathcal{S}_1 \subset V$ defined by

$$\mathcal{S}_1^l(z) = w^1 \otimes \ldots \otimes V_l \otimes \ldots \otimes w^d \tag{18}$$

Equation (17) is then equivalent, for a given $z = \otimes_{k=1}^{d} w^k \in \mathcal{S}_1$, to find $z^\diamond \in \mathcal{S}_1^l(z)$ such that

$$z^\diamond = \arg \min_{\hat{z} \in \mathcal{S}_1^l(z)} \mathcal{J}(u_m + \hat{z}) \tag{19}$$

Euler-Lagrange equation associated with problem (19) is

$$z^\diamond \in \mathcal{S}_1^l(z), \quad \mathcal{A}(u_m + z^\diamond, z^*) = \mathcal{L}(z^*) \quad \forall z^* \in \mathcal{S}_1^l(z) \tag{20}$$

Let us denote by $f_m^l(z)$ the unique solution of problem (20), where $f_m^l : \mathcal{S}_1 \to \mathcal{S}_1$ is a well-defined mapping. Starting from an initial guess $z^{(0)} \in \mathcal{S}_1$, we then construct a sequence $\{z^{(n)}\}_{n \in \mathbb{N}}$ defined by

$$z^{(n+1)} = f_m^d \circ \ldots \circ f_m^1(z^{(n)}) \tag{21}$$

In practice, we often observe a relatively fast convergence of the sequence $\{z^{(n)}\}$ towards an element $\check{z}$ which satisfies simultaneously $\check{z} = f_m^l(\check{z})$ for all $l \in \{1 \ldots d\}$. This behavior can be understood by the analogy between this alternated direction algorithm and a power iteration algorithm for capturing the dominant singular value [6].

3.2 Alternative definitions for general problems

3.2.1 Galerkin PGD

A direct application of PGD algorithm presented in section 3.1.3 is feasible, even for non-symmetric problems (2), with no associated convex optimization problem. Knowing $u_m \in \mathcal{S}_m$, we construct a sequence $\{z^{(n)}\}_{n \in \mathbb{N}}$ defined by iterations (21), where $f_m^l(z)$ is the solution of equation (20). $z^\diamond = f_m^l(z)$ is then a Galerkin approximation of $(u - u_m)$ on the subspace $\mathcal{S}_1^l(z)$. If the sequence $z^{(n)}$ converges towards an element $\check{z} \in \mathcal{S}_1$, it verifies simultaneously the Galerkin orthogonality conditions $\check{z} = f_m^l(\check{z})$, for all $l \in \{1, ..., d\}$. We then take $z_{m+1} = \check{z}$. We may also observe a non convergence of the sequence $z^{(n)}$. However, after a few iterations, the iterate $z^{(n)}$ can be selected as a good candidate for z_{m+1}.
Although there is no guaranty for convergence of the sequence u_m with m, good convergence properties of u_m are observed for a large class of linear problems [8, 15] or nonlinear problems [16], for the case of stochastic/deterministic separation, and for linear problems in the case of multidimensional separation [3]. However, for some nonsymmetric problems , the decomposition may present very bad convergence properties or even diverge [15]. This lack of robustness of Galerkin PGD has led to the proposition of other definitions of PGD, presented below.

3.2.2 Minimal Residual PGD

For non symmetric problems, a possible strategy consists in reformulating the problem as an optimization problem with a minimal residual formulation. Let us define the residual $R(u) = A(u) - l \in V$. The progressive Minimal Residual (MinRes) PGD is naturally defined by equation (13), with

$$\mathcal{J}(v) = \|R(v)\|^2 = (Au - l, Au - l) \tag{22}$$

An alternated minimization algorithm can then be applied in order to find a minimizer z_{m+1}. It consists in constructing the sequence $\{z^{(n)}\}_{n\in\mathbb{N}} \subset \mathcal{S}_1$ defined by (21), where for linear problems, mappings f_m^l are such that $z^\diamond = f_m^l(z)$ is the solution of

$$z^\diamond \in \mathcal{S}_1^l(z), \quad (A^*A(u_m + z^\diamond), z^*) = (A^*l, z^*) \quad \forall z^* \in \mathcal{S}_1^l(z) \tag{23}$$

where A^* is the adjoint operator of A. This construction of PGD is robust in the sense that the residual norm $\mathcal{J}(u_m)$ monotically decreases with m. However, these reformulations are often uneasy from theoretical and technical point of views. Moreover, they often lead to PGD decompositions with bad convergence properties with respect to usual norms [15], as it will be illustrated in the examples.

3.2.3 Minimax PGD

An alternative definition of PGD, called Minimax PGD, has been proposed in [15] in order to improve convergence properties of progressive PGD with respect to a specified metric. It can be interpreted as a Petrov-Galerkin PGD. It has been initially introduced for time-dependent partial differential equations. Here, it is extended to general non symmetric linear problems formulated in tensor product spaces. The construction of the Minimax PGD leads to computation times that are similar to those of the classical Galerkin PGD. Therefore, for a given accuracy, it usually leads to lower rank approximations and lower computation times, compared to Galerkin and Minimal Residual PGD.

Definition of the Minimax PGD. Let us introduce the functional $\mathcal{M}_m : V \times V \to \mathbb{R}$ defined by

$$\mathcal{M}_m(z, \tilde{z}) = \frac{1}{2}(z, z) - \mathcal{A}(u_m + z, \tilde{z}) + \mathcal{L}(\tilde{z}) \tag{24}$$

Let us note that inner product $(\cdot, \cdot)$ on V can be arbitrarily chosen. This choice has a consequence on the convergence properties of u_m. The progressive PGD u_m is then defined by

$$\mathcal{M}_m(z_{m+1}, \tilde{z}_{m+1}) = \max_{\tilde{z}\in\mathcal{S}_1} \min_{z\in\mathcal{S}_1} \mathcal{M}_m(z, \tilde{z}) \tag{25}$$

where two rank-one elements z_{m+1} and $\tilde{z}_{m+1}$ are constructed simultaneously. Let us note that the problem

$$\mathcal{M}_m(v^\diamond, \tilde{v}^\diamond) = \max_{\tilde{v}\in V} \min_{v\in V} \mathcal{M}_m(v, \tilde{v}) \tag{26}$$

admits the solutions $v^\diamond = u - u_m$, where u is the exact solution of the initial problem, and $\tilde{v}^\diamond = A^{*-1}v^\diamond$ (for A linear). z_{m+1} and $\tilde{z}_{m+1}$ can then be seen as rank-one approximations of $v^\diamond$ and $\tilde{v}^\diamond$ respectively.

Alternated direction algorithm. A possible way for constructing $(z_{m+1}, \tilde{z}_{m+1}) \in \mathcal{S}_1 \times \mathcal{S}_1$ consists in using an alternated direction algorithm, by solving successively problems of type

$$\min_{w^l\in V_l} \max_{\tilde{w}^l\in V_l} \mathcal{M}_m(\otimes_{k=1}^d w^k, \otimes_{k=1}^d \tilde{w}^k) \tag{27}$$

For given $(z, \tilde{z}) \in \mathcal{S}_1 \times \mathcal{S}_1$, problem (27) is equivalent to

$$\max_{\tilde{z}^\diamond\in\mathcal{S}_1^l(\tilde{z})} \min_{z^\diamond\in\mathcal{S}_1^l(z)} \mathcal{M}_m(z^\diamond, \tilde{z}^\diamond) \tag{28}$$

where linear spaces $\mathcal{S}_1^l(z)$ and $\mathcal{S}_1^l(\tilde{z})$ are defined by (18). For linear problems, the stationarity conditions associated with saddle point problem (28) are: find $(z^\diamond, \tilde{z}^\diamond) \in \mathcal{S}_1^l(z) \times \mathcal{S}_1^l(\tilde{z})$ such that

$$\mathcal{A}(u_m + z^\diamond, \tilde{z}^*) = \mathcal{L}(\tilde{z}^*) \quad \forall \tilde{z}^* \in \mathcal{S}_1^l(\tilde{z}) \tag{29}$$

$$\mathcal{A}(z^*, \tilde{z}^\diamond) = (z^*, z^\diamond) \quad \forall z^* \in \mathcal{S}_1^l(z) \tag{30}$$

Let us note that equations (29) and (30) can be solved one after the other. Equation (29) defines $z^\diamond$ as a Petrov-Galerkin approximation of $(u - u_m)$ in approximation space $\mathcal{S}_1^l(z)$, with test space $\mathcal{S}_1^l(\tilde{z})$. Equation (30) requires the solution of an adjoint problem. We denote by $(z^\diamond, \tilde{z}^\diamond) = f_m^l(z, \tilde{z})$ the solution of the system of equations (29)-(30). Starting for an initial guess $(z^{(0)}, \tilde{z}^{(0)}) \in \mathcal{S}_1$, we then define a sequence $\{(z^{(n)}, \tilde{z}^{(n)})\}_{n\in\mathbb{N}}$ defined by

$$(z^{(n+1)}, \tilde{z}^{(n+1)}) = f_m^d \circ \ldots \circ f_m^1(z^{(n)}, \tilde{z}^{(n)}) \tag{31}$$

In practice, we often observe a convergence of the sequence $\{(z^{(n)}, \tilde{z}^{(n)})\}_{n\in\mathbb{N}}$. However, this algorithm requires further theoretical investigations and possible improvements.

3.2.4 Use of traditional iterative solvers

An alternative way of addressing non symmetric or even nonlinear problems is to apply a standard global (nonlinear) iterative strategy for solving (2). Let us denote

by $\{u^{(n)}\}_{n\in\mathbb{N}}$ the sequence of iterates which are the solutions of the following linear problems:

$$u^{(n+1)} \in V, \quad \mathcal{B}_n(u^{(n+1)}, v) = \mathcal{B}_n(u^{(n)}, v) - (R(u^{(n)}), v) \quad \forall v \in V$$

where $\mathcal{B}_n$ is chosen as a symmetric continuous and coercive bilinear form on V. The PGD for linear symmetric problems presented in section (3.1.2) can then be applied for constructing a separated representation of $u^{(n+1)}$. From properties of $\mathcal{B}_n$, we have that the PGD strongly converges towards the exact iterate $u^{(n+1)}$. In fact, when introducing the PGD approximation, we only construct an approximate sequence $\hat{u}^{(n)}$, which can be seen as a perturbation of the original sequence $u^{(n)}$. The selection of a robust global iterative solver is then important. This solver must be robust with respect to perturbations in order to avoid a divergence of the sequence $\hat{u}^{(n)}$. This point will be investigated in a subsequent paper.

4 Updated Proper Generalized Decomposition

In many applications, a classical a posteriori tensor product approximation of the solution u reveals that a low rank approximation can be sufficient to obtain a good accuracy. However, in some cases, the progressive PGDs could lead to much higher rank approximations for the same accuracy, as it will be illustrated in the examples.

"Optimal" decompositions. In the case where problem (2) is equivalent to the optimization problem (12), a rank-m approximation $u_m \in \mathcal{S}_m$ could be naturally defined by

$$\mathcal{J}(u_m) = \min_{v_m \in \widetilde{\mathcal{S}}_m \subset \mathcal{S}_m} \mathcal{J}(v_m) \tag{32}$$

where $\widetilde{\mathcal{S}}_m$ is a suitable subset of $\mathcal{S}_m$ ensuring the existence of a minimizer depends (see [10, 12] for further discussions on the well-posedness of rank-m approximations in high dimension). In practice, a natural algorithm for computing such a minimizer still consists in using an alternated minimization procedure. However, the minimization along a direction l requires a minimization on $(V_l)^m$. Depending on the application, this problem may be unaffordable for large m. If the aim is to obtain the lowest rank approximation, this minimization problem should be tackled with. However, in the context of the solution of equations, we generally need an approximation with a good accuracy and not necessarily with the lowest rank.

Updates along selected directions. Even if we do not try to solve problem (32), updates along some selected directions may significantly improve the convergence of progressive PGDs. For a given element $u_m = \sum_{i=1}^m z_i \in \mathcal{S}_m$, with $z_i = \otimes_{k=1}^d w_i^k \in \mathcal{S}_1$,

and for a given direction $l \in \{1 \dots d\}$, we introduce the linear subspace $\mathcal{S}_m^l(u_m)$ defined by

$$\mathcal{S}_m^l(u_m) = \{\sum_{i=1}^{m} w_i^1 \otimes \dots \otimes \hat{w}_i^l \otimes \dots \otimes w_i^d \ ; \ \hat{w}_i^l \in V_l\}$$

For a given $u_m \in \mathcal{S}_m$, the following problem can be solved in order to improve the quality of the rank-m decomposition:

$$u_m^\diamond \in \mathcal{S}_m^l(u_m), \quad \mathcal{A}(u_m^\diamond, v_m) = \mathcal{L}(v_m) \quad \forall v_m \in \mathcal{S}_m^l(u_m) \tag{33}$$

It appears as an update of functions $\{w_i^l\}_{i=1}^m \in (V_l)^m$ in the decomposition u_m. When problem (2) corresponds to optimization problem (12), equation (33) is the Euler-Lagrange equation associated with the minimization of functional $\mathcal{J}$ on the linear subspace $\mathcal{S}_m^l(u_m)$. Even if there is no underlying optimization problem, the solution $u_m^\diamond$ of the update problem (33) is a Galerkin projection, which is well defined and usually allows to significantly improve the approximation u_m (see [8, 16, 15]).

5 Illustrations

5.1 Example 1

We consider the example of section (2.3.1) with $d = 2$ and $\Omega_1 = \Omega_2 = (0, 1)$. We take $\nu = 1/500$, $c = (1, 2)$, and $f = 1$. We introduce a finite element approximation space $V_h = V_{1,h} \otimes V_{2,h}$, where the $V_{l,h} \subset V_l = H_0^1(\Omega_l)$ are linear one-dimensional finite element spaces, with 500 nodes. We compute a reference Galerkin finite element approximation $u \in V_h$ of problem (2). Let us note that a stabilized finite element formulation could be adopted for this advection dominant equation. However, since we use a sufficiently fine mesh along each dimension, this stabilization is not required. The separation of spatial coordinates allows the introduction of very fine one-dimensional meshes and therefore, for some classes of problems, the PGD allows to circumvent the question of stabilization. We introduce the classical inner product in $L^2(\Omega)$, defined by $(u, v) = \int_\Omega uv \, dx$, and the associated norm $\|\cdot\|$. For estimating the convergence of PGD, we compute the error indicator $\epsilon_m = \|u - u_m\|/\|u\|$, where u is the reference Galerkin finite element approximation. This inner product is also used in the Minimax PGD for the definition of functional $\mathcal{M}_m$ in equation (24). A classical SVD of the reference solution is also computed. This SVD is defined with respect to the above inner product $(\cdot, \cdot)$. In dimension $d = 2$, the rank-m SVD $u_m \in \mathcal{S}_m$ is defined by $\|u - u_m\| = \min_{v_m \in \mathcal{S}_m} \|u - v_m\|$.

Convergence of PGDs. On figure 1, we plot the reference solution u and approximations u_m obtained with the three different progressive PGDs, for different ranks m. The three PGDs capture relatively quickly the main features of the solution but we clearly observe the superiority of the Minimax PGD in the present application.

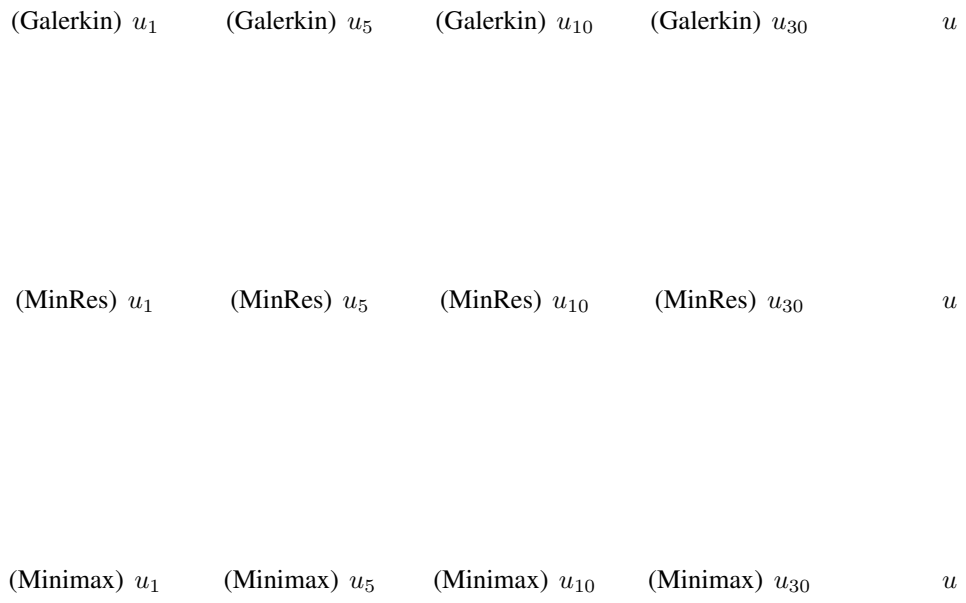

Figure 1: Convergence of progressive PGDs u_m (contour plots), compared with reference solution u.

Figure 2 shows the convergence curves of the different PGDs. On figure 2(a), we observe that progressive Galerkin and MinRes PGD have similar and relatively slow convergences, while the progressive Minimax PGD presents a much faster convergence. On Figure 2(b), we observe that updated PGDs (and especially Galerkin and Minimax PGDs) generate a decomposition which is very close to the SVD. However, as we will see below, the updating step significantly increases the computation time of the PGD in this application.

Convergence of the alternated direction algorithm. Figures 3 and 4 show the convergence of sequence $z^{(n)}$ generated by the alternated direction algorithm, for the computation of successive rank-one elements $z_i \in \mathcal{S}_1$, respectively for progressive PGDs (figure 3) and updated progressive PGDs (figure 4). The sequences $z^{(n)}$ are defined by equation (21) for the Galerkin and MinRes PGDs and by equation (31) for the Minimax PGD. On Figure 3, we observe a fast convergence of alternated direction algorithm for Galerkin and Minimax PGDs, for all the modes of the decomposition. For Galerkin (resp. Minimax) PGD, only $n = 3$ (resp. $n = 2$) iterations are sufficient in order to obtain a convergence with a precision of 10^{-2} for the computation of each modes z_i. However, we notice that for MinRes PGD, the sequence converges very slowly. The same conclusions are made on Figure 4 for the updated versions of

(a) Progressive PGDs (b) Updated PGDs

Figure 2: Convergence of progressive (a) and updated (b) PGDs. Reference SVD is also plotted.

PGDs. However, we notice that the update improves the convergence of the alternated direction for subsequent modes.

(a) Galerkin PGD (b) MinRes PGD (c) Minimax PGD

Figure 3: Convergence of the alternated direction algorithm for progressive PGDs. Stagnation error $\|z^{(n+1)} - z^{(n)}\|/\|z^{(n)}\|$ is plotted for $n = 1 \dots 5$ (the different curves) and for modes 1 to 60 (abscissa).

Computation times. The construction of the different PGDs leads to different computation times. On Figure 5 is plotted the error ϵ_m with respect to computation time for the different PGDs. We already noticed on figure 2 that the updating step (only one update along each direction) can significantly improve the convergence with m for all PGDs. However, this updating step significantly increases computation times and therefore, updated PGDs are not necessarily the fastest solution techniques. In this application, it appears that the progressive Minimax PGD (without update) is the fastest technique. This optimality is due to the relatively fast convergence of sequence u_m and the fast computation of alternated direction iterations. Updated Galerkin and Minimax PGDs lead also to quite good results. We note that MinRes PGD leads to higher computation times, although it leads to a good convergence of u_m. This is due to the fact that operator A^*A and right-hand sides $A^*(b - Au_m)$ have much higher ranks than operator A and right-hand sides $b - Au_m$.

(a) Updated Galerkin PGD (b) Updated MinRes PGD (c) Updated Minimax PGD

Figure 4: Convergence of the alternated direction algorithm for updated progressive PGDs. Stagnation error $\|z^{(n+1)} - z^{(n)}\|/\|z^{(n)}\|$ is plotted for $n = 1 \dots 5$ (the different curves) and for modes 1 to 60 (abscissa).

Figure 5: Error versus computation time for progressive and updated PGDs

5.2 Example 2

We consider the example of section (2.3.2). We let $d = 2$, with $\Omega_1 = \Omega_2 = (0, 1)$, and consider periodic boundary conditions, with $V_l = H^1_{per}(\Omega_l)$ for $l = 1, 2$. We choose for the diffusion parameter $\nu = 10^{\xi_1}$, with $\Xi_1 = (0, 1)$ with uniform probability measure P_{ξ_1}. It corresponds to a log-uniform distribution on $(1, 10)$ for random variable ν. We choose for the advection parameter $c = 125(\xi_2, 1)$, with $\Xi_2 = (0, 1)$ with uniform probability measure P_{ξ_2}. We choose $\sigma = 10$ and $f(x, y) = \exp(-300(x_1 - 1/4)^2) \exp(-300(x_2 - 1/2)^2)$. We introduce continuous linear finite elements for approximation spaces $V_{l,h} \subset V_l$, for $l = 1, 2$ (with 70 elements in each dimension), and finite elements with polynomial degree 4 for approximation spaces $V_{l,h} \subset V_l$, for $l = 3, 4$ (14 elements in each dimension). The dimension of approximation space $V_h = \otimes_{l=1}^4 V_{l,h}$ is $70^4 \approx 2.4\,10^7$. We introduce a classical inner product in $L^2(\Omega \times \Xi)$, defined by $(u, v) = \int_{\Omega\times\Xi} uv\, d\mu$, and the associated norm $\|\cdot\|$. For estimating the convergence of PGDs, we compute the error indicator $\epsilon_m = \|u - u_m\|/\|u\|$, where u is a reference solution, which is a fine approximation of the Galerkin approximation of problem (2) in V_h. The above inner product is also used in the Minimax PGD for the definition of functional $\mathcal{M}_m$ in equation (24).

Convergence of PGDs. Figure 6 illustrates the convergence curves of the different progressive and updated PGDs. It also illustrates the convergence of multidimensional singular value decompositions (MSVD). The progressive MSVD is defined by $u_m = \sum_{i=1}^m z_i$, with $\|u - u_{m+1}\| = \min_{z\in\mathcal{S}_1} \|u - u_m - z\|$. An alternated minimization algorithm is used for computing each $z_i \in \mathcal{S}_1$. The updated progressive MSVD is defined in the same way, with additional updates $\min_{v_m\in\mathcal{S}^l_m(u_m)} \|u - v_m\|$ along each direction l. Let us note that in dimension $d = 2$, progressive and updated MSVDs coincide, and coincide with a classical SVD. On figure 6(a), for progressive PGDs, we observe that Galerkin PGD does not converge and that the progressive Minimal Residual PGD has very poor convergence properties. The Minimax PGD has a better behavior and leads to a decomposition which is closer to progressive MSVD. Figure 6(b) illustrates a much better behavior of updated variants of PGDs (and MSVD). Let us note that even for the MSVD, the updating step leads to a significant improvement of the convergence. The three definitions of updated PGDs lead to quite similar convergence properties. However, we observe that MinRes PGD gives higher errors for a given rank. Even if MinRes PGD seems relatively robust with respect to convergence, it seems inefficient from a computational point of view, as illustrated on Figure 7.

On Figure 8, we plot rank-100 approximations u_{100} obtained with the three different progressive PGDs, for three samples of the parameters ξ. Figure (9) (resp. (10)) shows the approximations u_{30} (resp. u_{50}) obtained with updated PGDs for the same three samples. The reference solutions correspond to classical Galerkin finite element approximations in $V_{1,h} \otimes V_{2,h}$ for the two-dimensional problems associated with given values of the parameters ξ. For this application, we clearly notice the importance of the updating step and the necessity of computing high-rank representations for obtaining a good accuracy.

(a) Progressive PGDs (b) Updated PGDs

Figure 6: Convergence of progressive (a) and updated (b) PGDs and MSVDs

Figure 7: Error versus computation time for progressive and updated PGDs

6 Conclusion

We have presented and tested several definitions of Proper Generalized Decompositions (PGDs) for the a priori construction of separated representations of the solution of variational problems defined in tensor product spaces. These methods constitute very promising tools for circumventing the curse of dimensionality when addressing the numerical simulation of high dimensional models in computational science and engineering. The PGD method is relatively well mastered from a theoretical point of view for particular classes of problems associated with convex optimization problems. However, even for this class of problems, it appears that the different definitions of PGDs may lead to very different behaviors in practical applications. For more general problems, as illustrated in this paper, some variants of PGDs may lead to relatively poor convergence properties or even to non convergence. Many theoretical investigations are still necessary for a better understanding of the different variants of PGD methods and the introduction of more efficient algorithms for their construction. For many problems, the proposed PGD variants may fail at constructing a convergent separated representation of the solution. That proves the necessity of proposing new definitions of PGDs and associated robust algorithms, and also ad-hoc error estimation techniques dedicated to PGD methods.

(Galerkin) (MinRes) (Minimax) (Reference)

Figure 8: Progressive PGDs u_{100} (contour plots) for 3 samples of parameters ξ (from top to bottom).

References

[1] P. Ladevèze, J. Passieux, D. Néron, "The LATIN multiscale computational method and the Proper Generalized Decomposition", *Computer Methods in Applied Mechanics and Engineering*, 199(21-22): 1287–1296, 2010, doi:10.1016/j.cma.2009.06.023.

[2] F. Chinesta, A. Ammar, E. Cueto, "Recent advances in the use of the Proper Generalized Decomposition for solving multidimensional models", *Archives of Computational Methods in Engineering*, In press, 2010.

[3] A. Nouy, "Proper Generalized Decompositions and separated representations for the numerical solution of high dimensional stochastic problems", *Archives of Computational Methods in Engineering*, 2010, In press.

[4] P. Ladevèze, *Nonlinear Computational Structural Mechanics - New Approaches and Non-Incremental Methods of Calculation*, Springer Verlag, 1999.

[5] A. Ammar, B. Mokdad, F. Chinesta, R. Keunings, "A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modelling of complex fluids", *Journal of Non-Newtonian Fluid Mechanics*, 139(3): 153–176, 2006.

(Galerkin) (MinRes) (Minimax) (Reference)

Figure 9: Updated Progressive PGDs u_{30} (contour plots) for 3 samples of parameters ξ (from top to bottom).

[6] A. Nouy, "A generalized spectral decomposition technique to solve a class of linear stochastic partial differential equations", *Computer Methods in Applied Mechanics and Engineering*, 196(45-48): 4521–4537, 2007.

[7] C.L. Bris, T. Lelievre, Y. Maday, "Results and questions on a nonlinear approximation approach for solving high-dimensional partial differential equations", *Constructive Approximation*, 30(3): 621–651, 2009.

[8] A. Nouy, "Generalized spectral decomposition method for solving stochastic finite element equations: invariant subspace problem and dedicated algorithms", *Computer Methods in Applied Mechanics and Engineering*, 197: 4718–4736, 2008.

[9] J. Chen, Y. Saad, "On The Tensor Svd And The Optimal Low Rank Orthogonal Approximation Of Tensors", *Siam Journal On Matrix Analysis And Applications*, 30(4): 1709–1734, 2008.

[10] V. de Silva, L.H. Lim, "Tensor Rank and Ill-posedness of the best low-rank approximation problem", *SIAM Journal of Matrix Analysis & Appl.*, 30(3): 1084–1127, 2008.

[11] T.G. Kolda, B.W. Bader, "Tensor Decompositions and Applications", *SIAM Review*, 51(3): 455–500, September 2009.

(Galerkin) (MinRes) (Minimax) (Reference)

Figure 10: Updated Progressive PGDs u_{50} (contour plots) for 3 samples of parameters ξ (from top to bottom).

[12] A. Uschmajew, "Well-posedness of convex maximization problems on Stiefel manifolds and orthogonal tensor product approximations", *Numerische Mathematik*, 2010, ISSN 0029-599X (Print) 0945-3245 (Online).

[13] A. Falco, A. Nouy, "A Proper Generalized Decomposition for the solution of elliptic problems in abstract form by using a functional Eckart-Young approach", *Journal of Mathematical Analysis and Applications*, 2010, submitted.

[14] A. Nouy, A. Falco, "Proper Generalized Decomposition for nonlinear convex problems in tensor product Banach spaces", *In preparation*.

[15] A. Nouy, "A priori model reduction through Proper Generalized Decomposition for solving time-dependent partial differential equations", *Computer Methods in Applied Mechanics and Engineering*, 2010.

[16] A. Nouy, O. Le Maître, "Generalized Spectral Decomposition Method for Stochastic Non Linear Problems", *Journal of Computational Physics*, 228(1): 202–235, 2009.